

Computerized Adaptive Testing in Corporate E-learning

Duška Milanov¹, Miodrag Ivković², Kemal Kačapor³

¹Faculty of Sciences, University of Novi Sad, Serbia

²Technical faculty Mihajlo Pupin, University of Novi Sad, Serbia

³School of Economics and Business, Sarajevo, University of Sarajevo, Bosnia and Herzegovina

¹dusanka.milanov@gmail.com; ²misa.ivkovic@gmail.com; ³kemal.kacapor@efsa.unsa.ba

Abstract

The goal of this paper is to present and analyse the potential of E-learning and computerized adaptive testing, as a part of it, within a corporate context. Over the years, technological growth led to creation of learning organizations and knowledge economy so the solution for constant and cost-effective training of staff had to be found. That is where E-learning proved to be one of the most efficient solutions in a business organization. In the paper, features and segments of corporate E-learning will be presented. Meanwhile, fundamentals of computerized adaptive testing, item response theory and algorithm procedure are going to be described. Finally, potential of computerized adaptive testing used in determining employees knowledge and skills development is going to be discussed.

Keywords

E-learning; Corporate Training; Corporate Learning; Computerized Testing; Adaptive Testing; CAT

Introduction

Globalization and technological changes evolved so much in the last two decades that they have created new global economy which brought new business rules for companies. This economy is knowledge-based and is primarily leaning on use of ideas more than physical abilities of workforce, as well as on smart use of technology rather than simple transformation of raw materials or exploitation of inexpensive workforce. New ways of developing and using knowledge are created due to ever-growing need for innovations.

The new economy caused organizations and companies to recruit elites and develop staff with high competency. Thus, they need to provide development of skills and constant education to their workforce in order to keep up with changes and achieve success and financial growth. With E-learning and modern

technologies as tools for enabling that goal, they can propel corporate training to higher level.

Training and education in corporate organizations has always been present, whether in form of learning about the job itself and ways of completing various tasks, or managing relationships with colleagues and influence of work environment on society in general (Faherty; 2003). The modern market demands and emerging technologies changed the situation nowadays, forcing employees to constantly learn and improve their skills, and at the same time, enabling them to learn wherever and whenever is most suitable for them.

Many organizations handling large number of employees in corporate environments, are forced to use computer-based testing or computer assisted assessment tools, which offer effective and fast testing, self-paced testing, immediate feedback, and accurate, objective, and fast scoring. Current learning theories lead towards student-centred and personalized learning, pointing out the possibilities for eliminating cheating, reducing the examinee's anxiety, challenging but not frustrating the examinees, as well as for immediate and continuous examinee guidance based on knowledge, proficiency, ability, and performance (Economides, Roupas; 2007). Thus, many organizations are leaning towards computer adaptive testing (CAT) tools whose characteristics will be described in this paper, as well as their benefits and possible drawbacks.

E-learning in Corporate Environment

In business, learning is the process by which people acquire new skills or knowledge for the purpose of enhancing their capability in further work to realize their values in the corporation. It can be said that

learning is a conscious attempt on the part of organizations to improve productivity, effectiveness and innovativeness to meet market requirements. In the corporate context, training is the way the instruction is conveyed: it supports learning that is the individual's internal way of processing information into knowledge as in (Rosenberg; 2001).

E-learning is defined in different ways by different authors, but it always encompasses main idea of learning experiences delivered or enabled by electronic technology. According to (Rosenberg; 2001), E-learning refers to the use of internet technologies to deliver a wide array of solutions that enhance knowledge and performance. It is based on three fundamental criteria: (1) E-learning is networked, which makes it capable of instant updating, storage/retrieval, distribution and sharing of instruction or information; (2) It is delivered to the end user via a computer using a standard internet technology; (3) It focuses on the learning solutions that go beyond the traditional paradigms of training.

TABLE 1 DIFFERENT APPROACHES TO E-LEARNING

Approach	Description
Web based training (WBT)	This is a browser based equivalent to Computer based training, delivered via an intranet or Internet. It's advantage is the ability to deliver dynamic data that can be regularly updated from central location
Web based training plus mentorship	In addition to web-based training, the instructor/mentor is available for support via e-mail, chat sessions or the telephone
Synchronous E-learning	Involves the instructor and learners being in the same virtual classroom at the same time. An instructor at a central location leads a group of more than one learner at remote locations via an Internet/Intranet connection. The instructor converses with learners via audio and video
Asynchronous E-learning	This approach includes WBT but also recordings of synchronous sessions, delivered via the Internet on demand. The recordings can be controlled by the learner and replayed at will. This approach has the benefit of being available anywhere and anytime

According to (Urdan, Wegen; 2000), the term E-learning refers to a wide set of applications and

processes, including computer-based learning, web-based learning, virtual classrooms and digital collaboration. It can be defined as the delivery of content via all electronic media, including the Internet, intranets, extranets, satellite broadcast, audio/video tape, interactive TV and CD-ROM, as shown in TABLE 1.

However, every company has to examine its current state and needs in order to evaluate if E-learning has the potential to enhance employee learning and development programs. From an organizational viewpoint, it is necessary to examine the impact of E-learning on an organization and also to ensure that adequate IT, operational and financial resources are available to back up an E-learning initiative.

Segments of Corporate E-Learning

From a corporate viewpoint, there are several factors that need to be examined before adopting an E-learning solution which can be decomposed into several key areas. According to (Henry; 2001), a comprehensive E-learning solution includes three key elements:

Content

Intellectual property available in several formats such as text, video, audio, animation and simulation content which can be presented through web-based training courses, online documents, supplementary printed documents, multimedia presentations, video, audio, virtual labs, simulations and animations and tutor-led training via Internet. It is important for this content to be:

- (1) Delivered timely and updated,
- (2) Accurate, focused on important aspects and in reasonable quantities, and
- (3) Learning materials have to be relevant and according to learners needs.

Technology

Technological infrastructure plays important role in corporate training and enables delivery of content and knowledge through different channels. Four technologies important for corporate E-learning are:

- (1) Computer-based training (CBT) – lectures are created by using authoring software which enables creation of multimedia mix of different content forms, available for terminate users on

their computer, standing alone training stations and across local area networks.

- (2) Web-based training (WBT) – Wide use of the Internet enables the transition from CBT to WBT and leads to new opportunities for the learners. Corporate education became location and time independent and content could be updated and always available.
- (3) Learning management systems (LMS) – is a software package that enables the management and delivery of learning content and resources to learners. Most LMS systems are web-based to facilitate "anytime, anywhere" access to learning content and administration. The LMS allows registration of learners, the delivery and tracking of e-learning courses and content, testing, and may also allow the management of instructor-led training classes. In the most comprehensive of LMSs, one may find tools such as competency management, skills-gap analysis, succession planning, certifications, virtual live classes, and resource allocation (venues, rooms, textbooks, instructors, etc.). Most systems allow for learner self-service, facilitating self-enrolment, and access to courses. In corporate environment, LMS contributes to control of learning materials as it has centralized control over it, and to diversity of content as it can contain courses from various providers.
- (4) Learning objects – are reusable software components or units of learning material that can be extracted from one course and integrated into another, with the possibility of modification if necessary. The content of these objects is described with metadata element (in Extensible Markup Language format). Thus, they can be defined as collections of information objects assembled using metadata to better fit learner's needs and preferences. They enable flexibility of materials, simple update and content management, customization and interoperability.

Services

Hosting of E-learning software can be done by vendors or Application Service Providers who house or maintain it and the client can have access to it over the Internet.

Computerized Adaptive Testing

In the classical "paper and pencil" format, tests or assessments are identical for all the examinees and usually have a predetermined duration and a fixed number of questions which are featured with various levels of difficulty. The examinee's mark is usually the sum of the scores obtained in each of the questions and is regarded as competence criterion. Science recognized long time ago that giving a test that is much too easy for the candidates can turn out to be a waste of time, and can lead to unwanted candidate behaviour such as careless mistakes or deliberately choosing incorrect answers that might be the answers to "trick questions". On the other hand, questions that are much too hard, can produce generally inaccurate test results, because candidates try to attempt to answer the questions, often by guessing. Finally, the mark doesn't give enough information about the student answers which could help to determine unusual responses patterns or gaming behaviour from the learner; features that are important especially for diagnostic testing. Despite these problems, in most of the E-learning platforms, the tests often have the traditional testing format and are implemented as a simple online "paper and pencil" version.

This issue is addressed with use of Computer adaptive testing (CAT) which is a special case of Computer based testing (CBT), based on interactive method for assessing the level of a test-taker's knowledge, proficiency, ability, or performance using questions tailored to the specific student. The CAT system selects questions from a pool of precalibrated items appropriate for the level of the specific student.

A CAT system tailors the test to the proficiency of the individual examinee meaning it adjusts the test by presenting easy questions to a low-proficiency examinee and difficult questions to a high-proficiency examinee. As an alternative to giving each examinee the same fixed test, CAT item selection adapts to the ability level of individual examinees.

After each response the ability estimate is updated and the next item is selected such that it has optimal properties according to the new estimate (Triantafyllou, Georgiadou, Economides; 2008). The CAT first presents an item of moderate difficulty in order to initially assess each individual's level. During the test, each answer is scored immediately and if the examinee answers correctly, then the test statistically estimates examinee's ability as higher and then

presents an item that matches this higher ability. If the next item is again answered correctly, it re-estimates the ability as higher still and presents the next item to match the new ability estimate. The opposite occurs if the item is answered incorrectly. The computer continuously re-evaluates the ability of the examinee until the accuracy of the estimate reaches a statistically acceptable level or when some limit is reached; such as a maximum number of test items.

The overall score is determined by the level of the difficulty, and as a result, while all examinees may answer the same percentage of questions correctly, the high-ability ones will get a better score as they correctly answer more difficult items. Overall score of each examinee depends not only on the percentage of questions answered correctly but also on the difficulty level of these questions. Even if both examinees answer the same percentage of questions correctly, the high-proficiency examinee gets a higher score because the examinee answers correctly more difficult questions. Because each test is tailored to the individual examinee, far more information is gained from the examinee's response to each item than that in conventional test.

Item Response Theory

CAT require defining computational models which allow objective estimation and comparison of the proficiencies of learners that received different questions during a test. The psychometric technology that allows equitable scores to be computed across different sets of items is item response theory (IRT) which is also the preferred methodology for selecting optimal items typically selected on the basis of information rather than difficulty.

IRT is a statistical framework in which examinees can be described by a set of one or more ability scores that are predictive, through mathematical models, linking actual performance to test items, item statistics, and examinee abilities. It is a set of related psychometric models that provides a foundation for scaling persons and items based on responses to assessment items (Wainer, H. 2000). The person parameter usually is the proficiency or cognitive ability within a specific domain and it is represented by the Greek letter θ . A question in a test is a simple example of an item.

Much of the literature on IRT focuses on its models. Those models are usually functions relating person parameter and item parameters to the probability of a discrete outcome, such as a correct response to that

item. The three parameters logistic (3PL) model, first described by Birnbaum (1968), is the most widely used because it gives enough flexibility for implementation. Under the 3 parameter IRT model, the probability of a correct response to a given item i is a function of an examinee's true ability and three item parameters. Those three item parameters are the discrimination a , the difficulty b and the pseudo-guessing c . The later represents the chance for a low level examinee to find by guessing the correct response to the item. The conditional probability for a person with an ability θ to get a correct response to an item i (a_i , b_i and c_i) is given by the equation below.

$$p_i(\theta) = c_i + \frac{1 - c_i}{1 + e^{-a_i(\theta - b_i)}}$$

IRT provides strategies for estimating:

1. The a_i , or item discrimination, parameter,
2. The b_i , or item difficulty, parameter, and
3. The c_i , or guessing, parameter.

IRT has been successfully used in some computer adaptive testing system (Baker, 2001). IRT is a data oriented model and different studies have showed its advantages in terms of implementation complexity, predictive power and independence from a subjective expert appreciation.

CAT Algorithm

The CAT procedure is very simple and can be explained as: a test-taker is estimated (or guessed) to have a certain ability. An item of the equivalent level of difficulty is asked. If the test-taker succeeds on the item, the ability estimate is raised. If the test-taker fails in the item, the ability estimate is lowered. Another item is asked, targeted on the revised ability estimate and the process repeats.

Computer adaptive testing can begin when an item bank (or question pool) exists with IRT item statistics available for all items, when a procedure has been selected for obtaining ability estimates based upon candidate item performance, and when there is an algorithm chosen for sequencing the set of test items to be administered to candidates.

The CAT algorithm is usually an iterative process with the following steps:

1. All the items that have not yet been administered are evaluated to determine which will be the best one to administer next given the currently estimated ability level.

2. The "best" next item is administered and the examinee responds.
3. A new ability estimate is computed based on the responses to all of the administered items.
4. Steps 1 through 3 are repeated until a stopping criterion is met.

There are five technical components in building a CAT (Weis, Kingsbury; 1984):

1. Calibrated item pool - A pool of items (questions) must be available for the CAT to choose from. The pool must be calibrated with a psychometric model, usually IRT model.
2. Starting point or entry level - If some previous information regarding the examinee is known, it can be used, but often the CAT just assumes that the examinee is of average ability - hence the first item often being of medium difficulty.
3. Item selection algorithm - If the CAT has an estimate of examinee ability, it is able to select an item that is the most appropriate for that estimate which is done by selecting the item with the greatest information at that point.
4. Scoring procedure - After an item is administered, the CAT updates its estimate of the examinee's ability level. If the examinee answered the item correctly, the CAT will likely estimate their ability to be somewhat higher, and vice versa. This is done by using the item response function from item response theory to obtain a likelihood function of the examinee's ability.
5. Termination criterion - The decision as to when to stop a CAT test is the most crucial element. If the test is too short, then the ability estimate may be inaccurate. If the test is too long, then time and resources are wasted, and the items are exposed unnecessarily. The test-taker also may tire, and drop in performance level, leading to invalid test results.

The CAT test stops when:

1. The item bank is exhaustive.
2. The maximum test length is reached.
3. The ability measure is estimated with sufficient precision.
4. The ability measure is far enough away from the pass-fail criterion.
5. The test-taker is exhibiting off-test behaviour.

The CAT test cannot stop before:

1. A minimum number of items has been given.
2. Every test topic area has been covered.
3. Sufficient items have been administered to maintain test validity under challenge or review.

Use of CAT in Corporate Environment

Most organizations have numerous challenges when hiring new candidates, conducting job skills assessments as well as estimating future skills needs. Questions they face during these processes about improvement on the quality of employed candidates; areas of training which are really needed across the enterprise; or the identification of the individuals who need training within the organization; and ensuring the training programs are indeed effective.

Incorporating a CAT based assessment solution into daily recruitment process, skill audits, project staffing and training can enable companies to unleash the full potential of employee base and contribute to competitive advantage. Current solutions offer assessments on the conduction of employment screening using pre-hire testing, or identification of training requirements with pre-training assessments; even on the training effectiveness and knowledge transfer of the post-training assessment to provide career development with employment testing and certification programs.

Placing the person with the right skills and knowledge in the right position can significantly increase productivity and reduce staff turnover. Human resource experts estimate the costs associated with correcting hiring errors range between 1 to 2 times of the candidate's annual salary. Using a knowledge measurement solution to establish candidate suitability, organizations can reduce the administration time associated with hiring, increase the efficiency of the hiring process itself and improve the quality of candidate selection.

Most organizations execute generic training programs for entire departments without truly understanding what areas of training are required, and by whom. After much time and investment in training program execution, there are typically no mechanisms in place to determine overall training effectiveness and return on investment (ROI). Using a knowledge measurement solution to establish and evaluate training programs, companies could be able to determine whom in the organization needs training;

understand what areas of training should be focused on; optimize investment balance between training and organizational skill needs etc.

Advantages and Barriers of CAT Usage

In general, computerized testing greatly increases the flexibility of test management. Some of the benefits are:

- Tests are given "on demand" and scores are available immediately.
- Neither answer sheets nor trained test administrators are needed. Test administrator differences are eliminated as a factor in measurement error.
- Tests are individually paced so that a examinee does not have to wait for others to finish before going on to the next section. Self-paced administration also offers extra time for examinees who need it, potentially reducing one source of test anxiety.
- Test security may be increased because hard copy test booklets are never compromised.
- Computerized testing offers a number of options for timing and formatting. Therefore it has the potential to accommodate a wider range of item types.
- Significantly less time is needed to administer CATs than fixed-item tests since fewer items are needed to achieve acceptable accuracy. CATs can reduce testing time by more than 50% while maintaining the same level of reliability. Shorter testing times also reduce fatigue, a factor that can significantly affect an examinee's test results.
- CATs can provide accurate scores over a wide range of abilities while traditional tests are usually most accurate for average examinees.

In terms of corporate benefits, CAT can help achieve the following: increase employee competence (improve employee productivity, hiring and selection process, reduce employee attrition, increase customer credibility); increase job satisfaction (reduce training time and cost, improve training ROI, provide objective ROI confirmation); reduce hiring time and costs (reduce staff attrition, use objective benchmarks to differentiate candidates).

The first barrier encountered in CAT is the calibration of the item pool (Weiner, 2000.). In order to model the characteristics of the items (e.g., to pick the optimal

item), all the items of the test must be pre-administered to a sizable sample and then analysed. To achieve this, new items must be mixed into the operational items of an exam (the responses are recorded but do not contribute to the test-takers' scores), called "pilot testing" or "pre-testing". This presents logistical, ethical, and security issues. For example, it is impossible to field an operational adaptive test with brand-new, unseen items; all items must be pretested with a large enough sample to obtain stable item statistics.

Some other CAT limitations also raise several technical and procedural issues:

- CATs are not applicable to all subjects and skills. Most CATs are based on an item-response theory model, yet item response theory is not applicable to all skills and item types.
- Hardware limitations may restrict the types of items that can be administered by computer. Items involving detailed art work and graphs or extensive reading passages, for example, may be hard to present.
- CATs require careful item calibration. The item parameters used in a paper and pencil testing may not hold with a computer adaptive test.
- CATs are only manageable if a facility has enough computers for a large number of examinees and the examinees are at least partially computer-literate. This can be a big limitation.
- The test administration procedures are different. This may cause problems for some examinees.
- With each examinee receiving a different set of questions, there can be perceived inequities.
- Examinees are not usually permitted to go back and change answers.

Conclusions

In this time of global market competition, rapid technological advances, demographic changes, and service/knowledge-based economy force, organizations need to educate and constantly train their workforce. Global use of the Internet allows organizations to have new possibilities for delivering and sharing knowledge and education resources. It also brings current, dynamic educational content, and personalized, relevant learning experiences, with

ongoing collaboration and communication with experts and mentors. Current main obstacles to e-learning, such as bandwidth, content availability, technology standards, and service integration, are constantly overcome and are not considered being big problem areas much longer.

Although it has taken some time for the business world to understand the benefits of E-learning, it is clear that limitless opportunities are ahead. Employing a successful e-learning strategy allows a corporation to cut costs significantly, while increasing workplace satisfaction and raising employee motivation.

Adaptive tests are designed to maximize measurement efficiency, or the precision of test scores in relation to test length. This means an adaptive test can either save time shorter than a conventional test with equal precision, or improve score quality more significant than a conventional test of equal length. The examinees with the most benefit to gain are those at either the high or low extremes of the performance continuum. They are usually poorly served by conventional tests, which are generally designed to best fit the average test-taker.

ACKNOWLEDGMENT

This research is supported by Ministry of Science and Technological Development, Republic of Serbia, under the project number TR32044 "The development of software tools for business process analysis and improvement".

REFERENCES

- Baker, Francis. The Basics of Item Response Theory. ERIC Clearinghouse on Assessment and Evaluation, University of Maryland, 2001.
- Economides, Anastasios A. and Chrysostomos Roupas. "Evaluation of Computer Adaptive Testing Systems." *International Journal of Web Web-Based Learning and Teaching Technologies*, Vol. 2, Issue 1 (2007) : 70-87.
- Faherty, Rodger. "Corporate E-learning." Dublin Institute of Technology, School of Computing Research Paper, 2003.
- Henry, Paul. "E-learning Technology, Content and Services." *Education & Training*, Vol. 43, 2001.
- Linden, Wim J. van der and Cees A.W. Glass. *Computerized Adaptive Testing: Theory and Practice*. Kluwer Academic Publishers, 2003.
- Liu, Leping, D. LaMont Johnson, Cleborne D. Maddux and Norma J. Henderson. *Evaluation and Assessment in Educational Information Technology*, The Haworth Press, Inc, 2001.
- Lynch, Grace and Rhys James. "Staying Relevant and Current with Online Learning in an Increasingly Global and Competitive Environment." *International Journal of Advanced Corporate Learning*, Vol. 5, No.3 (2012): 26-28.
- Makransky, Guido. "Computerized Adaptive Testing in Industrial and Organizational Psychology." University of Twente, 2012.
- Urdan, Trace A. and Cornelia C. Weggen. *Corporate E-learning: Exploring a New Frontier*, WRHambrecht + Co, 2000.
- Rosenberg, Mark J. *E-Learning – Strategies for Delivering Knowledge in the Digital Age*, The McGraw Hill Companies, Inc, 2001.
- Sitnikov, Sergey, Boris Kruk, Olga Zhuravleva and Natalia Chupakhina. "Corporate E-learning Strategy." *International Journal of Advanced Corporate Learning*, Vol. 3, No.4 (2010): 41-44.
- Triantafillou, Evangelos, Elissavet Georgiadou and Anastasios A. Economides. "The Design and Evaluation of a Computerized Adaptive Test on Mobile Devices." *Computers & Education* 50 (2008): 1319-1330.
- Wainer, Howard. *Computerized Adaptive Testing: A primer*. Mahwah, NJ: Lawrence Erlbaum Associates, 2000.
- Weiss, David J. and Cage G. Kingsbury. *Application of Computerized Adaptive Testing to Educational Problems*. *Journal of Educational Measurement*, 21, 1984: 361-375.